

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,500

Open access books available

136,000

International authors and editors

170M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# The Multi-Tier Instrument in the Area of Chemistry and Science

*Habiddin Habiddin and Septiana Ayuningrum Nofinadya*

## Abstract

Knowledge of students' unscientific understanding before learning a new topic known as students' preconception or prior knowledge is vital for helping the teacher design a proper teaching strategy. Meanwhile, knowledge of students' understanding after teaching will provide a way for a teacher to evaluate the effectiveness of his/her teaching. For these reasons, science educators should investigate students' understanding over time. Studying students' understanding requires a proper and powerful tool/instrument such as a multi-tier instrument. This paper describes the history of multi-tier instruments initiated by the two-tier and recently became a five-tier instrument, the procedure to develop the instrument, and how to utilize the instrument to identify students' unscientific understanding. Our recent study describing the development of a four-tier instrument of electrolyte and non-electrolyte solution (FTI-ENES) is presented.

**Keywords:** multi-tier instrument, four-tier instrument, three-tier instrument, two-tier instrument, five-tier instrument, unscientific understanding, misconception, science assessment

## 1. Introduction

Investigating students' in-depth understanding, mainly their unscientific knowledge, has been carried out for decades. Teachers' knowledge of students' understanding, including their prior knowledge or preconception and understanding after teaching, is valuable. Knowledge regarding students' preconceptions is essential in assisting educators in providing effective teaching and learning. Many studies have proved the contribution of students' prior knowledge to their teaching success [1, 2]. Several instruments have been used for uncovering students' conception in science, including concept mapping [3], interviews [4], and the multiple-choice test [5, 6]. A proper and effective instrument must be utilized to investigate students' understanding. A typical instrument such as a multiple-choice question (MCQ) cannot uncover a deep understanding [7] in science, particularly students' unscientific understanding/misconceptions. It has been revealed that the previous instruments have some disadvantages. Concept mapping relies on students' ability to master vocabulary [8], while the interview is time-consuming [9]. For multiple-choice questions, students' test-wiseness skills [10] could affect their reliability and validity indices, and the reason for students' answers cannot be fully uncovered [11]. Also, the role of guessing is often dominant in a multiple choice question [12].

Due to those previous instruments' disadvantages, the multi-tier format's diagnostic tool has recently been one of the most frequent instruments applied

in science education studies. Our previous study [13] investigated the instrument used in the study involving students' understanding of chemistry and other science disciplines (biology and physics) covered in Indonesian journals. We revealed that multi-tier instruments, particularly four-tier instruments, have been the most accepted instrument and widely applied by Indonesian researchers in identifying students' unscientific understanding.

In this paper, several terminologies, including students' conception, students' understanding, students' scientific understanding, students' scientific knowledge, students' unscientific understanding, and misconceptions, are found. Students' conception reflects students' ideas and mental processes regarding natural phenomena. The ideas could be relevant or irrelevant to the concept accepted by the scientific community [14]. For this reason, the terminology of students' conception and students' understanding are interchangeable in this paper. The ideas which adhere to the concept accepted by the scientific community are called scientific knowledge. In contrast, those different from a view taken by the scientific community are called unscientific understanding.

The incorrect idea harbored by any particular person has been described in several different terminologies in the scientific literature, including wrong knowledge, misconception, erroneous ideas, unscientific understanding, alternative conception, misunderstanding, erroneous concepts, naïve idea, alternative frameworks, naïve concept, misinterpretation, and oversimplifications. Although these terms are interchangeable, the "unscientific understanding" is preferred in this paper because it reflects the nature of students' incorrect ideas or concepts.

## **2. The development of multi-tier instrument: the chronological perspective**

### **2.1 Two-tier instrument: The milestone of multi-tier instruments**

The use of multi-tier instruments in science education was initiated by Treagust [15], investigating students' unscientific understanding in particular. The example of the two-tier instrument applied in such an instrument's initial development is provided in **Figure 1**.

The first-tier at the initial format portrayed in **Figure 1** consists of a multiple-choice question (MCQ) with only two options (one correct answer and one incorrect answer). This MCQ with a two-options format is quite uncommonly applied in science assessment, common in at least four options. The second tier consists of four statements covering the reasons for students' answers to the first-tier. The four reasons consist of one valid or scientific reason and three wrong or unscientific reasons. The combination of students' incorrect answers and the incorrect reason is the basis for revealing students' unscientific understanding or misconception. All incorrect reasons in the reason tier are composed based on students' actual unscientific understanding obtained from preliminary tests, interviews, and literature. The next generation of the two-tier instrument has employed a more standard MCQ in the first-tier, as depicted in **Figure 2**.

This two-tier format has been applied to investigate students' conception in many science education research including Tan et al. [16] in inorganic chemistry, Tuysuz [17] in Separation of Matter, Griffard & Wandersee [18] in Photosynthesis, Chandrasegaran et al. [9] in Chemical reaction, Peterson et al. [5] in covalent bonding, Tyson et al. [19] in chemical equilibrium, Adadan & Savasci [20] in solution chemistry and many others.

Water (H<sub>2</sub>O) and hydrogen sulphide (H<sub>2</sub>S) have similar chemical formulae and have V-shaped structures. At room temperature, water is a liquid and hydrogen sulphide a gas. The difference in state between water and hydrogen sulphide is due to the presence of strong intermolecular forces between

(1) H<sub>2</sub>O molecules      (2) H<sub>2</sub>S molecules

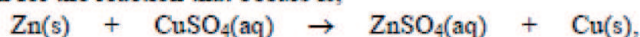
Reason

- (a) The difference in strength of the intermolecular forces is due to differences in the strength of the O–H and S–H covalent bonds.
- (b) The bonds in H<sub>2</sub>S are easily broken, whereas in H<sub>2</sub>O they are not.
- (c) The difference in strength of the intermolecular forces is due to the difference in polarity of the molecules.
- (d) The difference in strength of the intermolecular forces is due to the fact that H<sub>2</sub>O is a polar molecule, and H<sub>2</sub>S is a non-polar molecule.

**Figure 1.**  
 Example of the two-tier instrument developed by Treagust [15].

When powdered zinc is added to blue aqueous copper(II) sulfate and the mixture shaken, the blue colour of the solution gradually fades and it becomes colourless. At the same time a reddish-brown deposit is produced.

The chemical equation for the reaction that occurs is,



while the ionic equation is,  $\text{Zn(s)} + \text{Cu}^{2+}(\text{aq}) \rightarrow \text{Zn}^{2+}(\text{aq}) + \text{Cu(s)}.$

Why did the solution finally become colourless?

- A Copper has formed a precipitate.
- B Zinc is more reactive than copper(II) sulfate.
- C The copper(II) sulfate has completely reacted.
- D Zinc has dissolved, just like sugar dissolves in water.

The reason for my answer is:

- 1 Zinc ions are soluble in water.
- 2 Zinc loses electrons more readily than copper.
- 3 Soluble, blue Cu<sup>2+</sup> ions have formed insoluble, reddish-brown copper atoms.
- 4 In aqueous solution Cu<sup>2+</sup> ions produce a blue solution, while Zn<sup>2+</sup> ions produce a colourless solution.

**Figure 2.**  
 Example of the next generation of the two-tier instrument developed by Chandrasegaran et al. [9].

## 2.2 Three-tier instrument

After being applied in many studies, science education researchers realized that the two-tier instrument has deficiencies. Students selected the correct answer and correct reason randomly without holding a scientific reason to the relevant concept on certain occasions. The role of guessing and the actual unscientific understanding are difficult to be differentiated in a two-tier instrument [21, 22].

To overcome the two-tier instrument's drawback, a three-tier instrument was developed with the additional confidence rating tier, as shown in **Figure 3**. The third-tier requires students to state whether they are sure or unsure of their answer and reason. A correct answer and reason with a sure expression imply a scientific understanding. Meanwhile, an incorrect answer and reason with a sure expression imply an unscientific understanding or misconception. An incorrect answer and reason with an unsure expression imply that the incorrect answer is not a result of



The ozone layer,

- a. Protects the Earth from acid rain.
- b. Filters the ultraviolet (UV) rays of the sun.
- c. Helps to keep the Earth's temperature stable to make it livable.

Which one of the following is the reason for your answer to the previous question?

- a. The ozone layer absorbs the sun's ultraviolet (UV) rays which is potentially damaging to life on the Earth.
- b. The ozone layer prevents sun rays to exit from the atmosphere, consequently keeps it warm enough to live.
- c. The ozone layer works as a kind of shield, so it does not let acid rain to reach the Earth's surface.
- d. ....

Are you sure about your answer given to the previous two questions?


- a. Yes.
- b. No.

Figure 3.  
Example of a three-tier instrument developed by Arslan et al. [23].

misconception or unscientific understanding; instead, it lacks knowledge or guessing. This aspect distinguishes the three-tier format and the previous format. The same pattern of the three-tier instrument portrayed in **Figure 3** has been used in the following studies [11, 24, 25].

The subsequent development of a three-tier instrument utilized a more flexible confidence rating with a broader range of confidence, as displayed in **Figure 4**. This pattern seems to have been influenced by the standard confidence rating scales applied in many four-tier instruments that had been published before this three-tier work was carried out.

1.1



The system shown on the left figure represents a closed flask connected with a balloon and filled with air at 40 °C. If the temperature of the system was decreased from 40 °C to 5 °C, how would the volume of the balloon change? (Atmospheric pressure is 1 atm for both temperatures.)

- A) It would decrease.
- B) It would not change.
- C) It would increase.

1.2 Which one of the followings is the reason of your answer for the previous question?

- A) If the temperature was decreased, air particles in the system would shrink.
- B) If the temperature was decreased, the pressure in the system would decrease.
- C) If the temperature was decreased, the pressure in the system would increase.
- D) If the temperature was decreased, the distance between air particles would decrease.
- E) Atmospheric pressure is constant.
- F) .....

1.3 How sure you are about your answers for the previous two questions?

	Unconfident						Confident
Confidence Rating:	1	2	3	4	5	6	7

Figure 4.  
Example of a three-tier instrument developed by Aydeniz et al. [26].

2.3 Four-tier instruments

The confidence rating index (CRI), which is only attached to the third tier of the three-tier instrument, leads to an unclear message whether students have the same or different confidence levels between their answer and their reason [23]. For this reason, many science education researchers developed and applied the four-tier instrument. The first-tier, called Answer-tier (A-tier), consists of MCQ with several options (commonly 4). The second tier is the confidence rating for the A-tier. The third-tier, which is called Reason-tier (R-tier), consists of several statements with one correct statement relevant to the selected answer and several unscientific statements. The fourth-tier is the confidence rating for the R-tier.

The confidence rating index (CRI) for A-tier and R-tier ranged from 1 (just guessing) to 6 (absolutely confident). This more comprehensive range was then adopted for some studies that utilize three-tier instruments, as shown in **Figure 4**. In our recent works [7], we prefer to apply five scales of confidence rating instead of 6 scales (**Figure 5**).

Using five scales of CRI provides better clarity in differentiating students' level of confidence ratings. For example, the difference between 'confident' [4], 'very confident' [5], and 'absolutely confident' [6] in a six scales CRI format is quite challenging to be recognized. However, 'quite confident' [4] and 'very confident' [5] in 5 scales format is more comfortable to be understood. When a student is 100% sure of his/her answer, he/she will state very confident. Meanwhile, when he/she is not 100% sure of his/her answer, he/she will state quite confident. 'Average' [3] is used to express an equal portion of sure and unsure, which is not available in the six scales format. 'Very unconfident' [1] is used to express 100% unsure, including guessing or absolutely no knowledge regarding the concept. While 'not very confident' [2] is used to express an unsure reason with a small portion of feeling that his/her answer may be correct. For this reason, we suggest using five scales of CRI instead of 6 scales (**Figure 6**).

The current development of a multi-tier instrument is a five-tier instrument published by Anam et al. [28], with the additional fifth tier in which students are required to provide a draw/pictorial representation of his/her answer. This

Consider a real gas placed in a container. If intermolecular attractions were to disappear suddenly, which one of the following is likely to occur?

**Answer**

(a)

the pressure decreases

(b)

the pressure increases (\*)

(c)

the pressure remains the same

(d)

the gas expands

**Confidence Rating for Answer:**

1

2

3

4

5

6

Just Guessing

Very Unconfident

Unconfident

Confident

Very Confident

Absolutely Confident

**Reason**

(a)

the gas molecules are further away from each other

(b)

the frequency of collisions of the molecules with the walls of the container increases (\*)

(c)

the gas molecules will have more freedom to move around

(d)

It will behave more like an ideal gas,

**Confidence Rating for Reason:**

1

2

3

4

5

6

Just Guessing

Very Unconfident

Unconfident

Confident

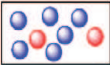
Very Confident

Absolutely Confident

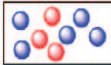
Figure 5.  
Example of four-tier instrument with six confidence ratings [27].

For a hypothetical reaction:  $X + Y \rightarrow \text{Products}$ , the rate of reaction is second order with respect to X but first order with respect to Y. Four experiments are carried out with different starting concentrations represented pictorially below in the boxes A, B, C and D. Which of the starting conditions (A, B, C or D) will result in the highest rate of reaction?

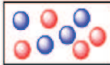
A



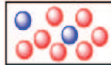
B



C



D



● = X  
● = Y

State the confidence rating of your answer  
 1. Very unconfident    2. Not very confident    3. Average    4. Quite confident    5. Very confident

Which one of the following options is the reason for your answer to the question?

A. It has the highest concentration of the reactant which is 2<sup>nd</sup> order  
 B. The concentrations of both reactants are the same, therefore the ratio of collision is more favourable  
 C. The amount of X and Y and the average of each keep determine the rate  
 D. The concentration of Y is much higher than the concentration of X and this leads to the reaction being completed faster

State the confidence rating of your answer  
 1. Very unconfident    2. Not very confident    3. Average    4. Quite confident    5. Very confident

**Figure 6.**  
Example of four-tier instrument in chemical kinetics with five confidence ratings [7].

additional drawing will ensure the mental model of the students can be uncovered. Even though the work in a five-tier instrument is still limited, we believe that it offers a more powerful tool in this regard. A pictorial tool is supported by psychology cognitive theory that helps students solve a multistep task [29].

### 3. The procedure in developing a multi-tier instrument

Treagust [15] proposed the two-tier instrument development is the fundamental development of the next generations of multi-tier instruments, including three-tier and four-tier instruments. Treagust [15] employed ten steps with three board categories in developing a two-tier instrument. The first four steps are named *defining the content*. Steps 5, 6, and 7 are named *obtaining information about students' misconceptions*. The last three steps are named as *developing a diagnostic test*. The steps are:

1. Identifying proportional knowledge statements
2. Developing a concept map
3. Relating proportional knowledge to the concept map
4. Validating the content
5. Examining related literature
6. Conducting unstructured student interviews
7. Developing multiple-choice questions with free responses
8. Developing the two-tier diagnostic tests
9. Designing a specification grid
10. Continuing refinements

When we developed a four-tier instrument in the area of chemical kinetics named FTDICK [7], we simplified the procedure to be six steps as the following. This procedure is applicable to developing multi-tier instruments.



### 3.1 Step 1: Mapping concept

In this step, several essential concepts in a particular topic are identified concerning the concept's scope in the relevant curriculum. For example, when we developed a four-tier instrument to identify secondary school students' understanding of thermochemistry, the competence mastery indicator document (Indikator Pencapaian Kompetensi, IPK) in the syllabus for Indonesian chemistry secondary school was considered. System and surrounding, enthalpy, exothermic reaction, and endothermic reaction are essential concepts in the Indonesian curriculum. When we developed a four-tier instrument of chemical kinetics for first-year chemistry students, university students' chemistry curriculum was considered. Rate law, the relation between reactant concentration and time, temperature and rate, activation energy, and reaction mechanisms are essential concepts for first-year university students.

### 3.2 Step 2: Developing the multiple-choice question with free responses (MCQ-FR)

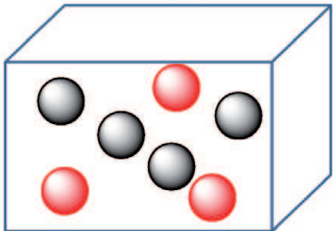
Each essential concept should be represented by two or more questions to ensure that it reflects all the competence and knowledge that should be mastered at the concept. **Figure 7** below depicts an example of MCQ-FR in the concept of chemical kinetics, particularly rate law and the relation of concentration and rate.

### 3.3 Step 3: Validating the MCQ-FR

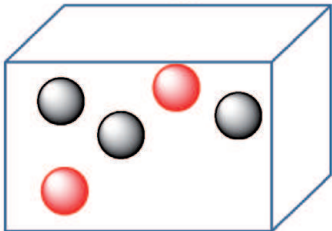
Before it is used to collect the preliminary data, the content of MCQ-FR, the relevance with curriculum, and language clarity are assessed to get feedback from some experts in the field. This feedback will be the basis to revise the MCQ-FR.

The experimentally determined rate law for the reaction  $\text{H}_2(\text{g}) + \text{I}_2(\text{g}) \rightarrow 2\text{HI}(\text{g})$  is  $\text{rate} = k [\text{H}_2] [\text{I}_2]$ . Each of boxes below represents a reaction mixture in which  $\text{H}_2$  molecule is shown as red spheres and  $\text{I}_2$  molecule is the black one.

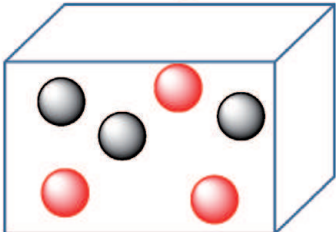
**A**



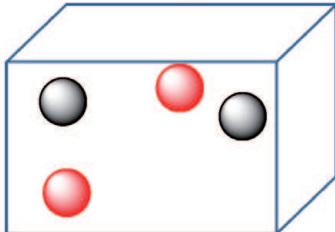
**B**



**C**



**D**



Arrange these mixtures based on the increasing rate of reaction!

**State your reason here:**

.....

**Figure 7.**  
 Example of MCQ-FR in chemical kinetics.



### 3.4 Step 4: Testing and collecting students' unscientific understanding

The revised MCQ-FR is then used to collect preliminary data, which are students' unscientific understanding or illogical reasons. For example, in answering the question in **Figure 7**, some students believed that *option D would be the highest rate because the concentration of two reactants ( $H_2$  and  $I_2$ ) is the same*. These illogical reasons are then collected and employed as the basis to develop the prototype multi-tier instrument.

### 3.5 Step 5: Developing the prototype multi-tier instrument

A significant number of students should demonstrate students' unscientific understanding used as a reason option. Students' responses in this step are also used to measure the MCQ-FR quality in terms of validity, reliability, distractor effectiveness, discriminatory index, and difficulty level. The unscientific understanding above is utilized as the optional reason at the multi-tier instrument (**Figure 6**, Reason B).

### 3.6 Step 6: Validating the prototype and refining the final multi-tier instrument

The next step is testing the prototype multi-tier instruments to a group of students to measure its validity, reliability, distractor effectiveness, discriminatory index, and difficulty level (5 parameters). This step is also named empirical validity. Please refers to the educational evaluation and measurement references to find out the formulae to calculate these parameters. The analysis of the five parameters' values is the basis for revising the prototype and producing the final multi-tier instrument, which applies to the broader community.

## 4. Grading students' responses and how to determine students' unscientific understanding level

### 4.1 Treatment of data

Students' responses to the multi-tier questions provide four types of combinations of students' answers and reasons, namely: Correct Answer and Correct Reason (CACR) representing good scientific understanding; Correct Answer and Wrong Reason (CAWR) representing a false positive of students' unscientific understanding; Wrong Answer and Correct Reason (WACR) representing a false negative of students' unscientific understanding. These three categories are not discussed widely in this paper. Wrong Answer and Wrong Reason (WAWR) represents an actual student's unscientific understanding. This WAWR is the central aspect discussed in this regard and the prime category to be used in interpreting students' unscientific understanding.

### 4.2 Parameters to classify students' unscientific understanding

Students' unscientific understanding is determined based on students' WAWR combinations. Several parameters and terminologies have been used to determine the level of students' unscientific understanding based on the students' confidence ratings or confidence rating index (CRI) of WAWR. Caleon & Subramaniam [21] employed six scales of confidence ratings and classified unscientific understanding or misconception as to the following. A *genuine* unscientific understanding is an unscientific understanding expressed with a  $CRI \geq 3.5$ . Meanwhile, a *spurious*

CRI	Category
$\geq 3$	Genuine unscientific understanding
	3–4: Moderate unscientific understanding
	$\geq 4$ : Strong unscientific understanding
$< 3$	Spurious unscientific understanding

**Table 1.**  
The parameter to classify unscientific understanding for 5 CRI scales.

unscientific understanding is an unscientific understanding expressed with a CRI  $< 3.5$ . *Genuine* unscientific understanding is further categorized into *moderate* unscientific understanding (those expressed with medium level CRI - between 3.5 and 4.0) and high level of unscientific understanding (those expressed with a high CRI of 4.0 and above). Literature using this scale [1–6] considers 3.5, i.e., the mid-point of unconfident and confidence as the limit of a *genuine* misconception.

The use of this parameter with a decimal number (3.5 as the limit) raises a critique considering that all the CRI scales are in whole numbers. Therefore, the rationale to use the decimal limit is questionable. For this reason, we suggest using the following parameter to classify students’ unscientific understanding for a multi-tier instrument that employs five scales of CRI (**Table 1**).

The example of how to determine students’ unscientific understanding is provided from our work in the area of thermochemistry, which is in the press for publication elsewhere. The question in **Figure 8** was intended to investigate students’ understanding of the system and surroundings, particularly the difference between open, closed, and isolated systems.

In answering the question in **Figure 8** above, 34.43% of students demonstrated an unscientific understanding that the *drop of water in the bottle’s outer wall comes from the bottle’s melting ice*. This unscientific understanding was demonstrated by those provided WAWR combination and also CAWR combination. The WAWR combination was with answer A - Reason B, while the CAWR combination was mostly with Answer B - Reason B. To justify that the unscientific understanding is *genuine* or *spurious*, the CRI must be taken into account. If the CRI of whom provided WAWR and/or CAWR combinations is 4.0, it can be declared that the unscientific understanding is *genuine* and fall in the *moderate* category. If the CRI of those provided WAWR and/or CAWR combinations is 3.0, it can be declared that the unscientific understanding is *spurious* and is a result of a *lack of knowledge* rather than a misconception.

An ice cube is placed in a closed plastic bottle. After sometime, the ice melts and some drops of water are observed at the outer wall of the bottle. Referring to this phenomenon, it can be concluded that it is a/an....

A. open system

B. closed system

C. isolated system

Please state your confidence rating for the answer

1. very unconfident/guessing

2. unconfident

3. moderate

4. confident

5. very confident

State the reason for your answer

A. There is energy exchange between system and surrounding but not for the matter exchange

B. Drops of water confirm the matter exchange between system and surrounding

C. The is no exchange of matter and energy between system and surrounding

Please state your confidence rating for the answer

1. very unconfident/guessing

2. unconfident

3. moderate

4. confident

5. very confident

**Figure 8.**  
Example of a four-tier instrument in thermochemistry [30].

## 5. Development of four-tier instrument in the topic of electrolyte and non-electrolyte solution (FTI-ENES): an empirical study

This section will present our current study in this area involving the development of a four-tier instrument in the topic of electrolyte and non-electrolyte solution. The instrument that was produced in this study is named the Four-Tier Instrument of Electrolyte and Non-Electrolyte Solution (FTI-ENES).

### 5.1 Method

This research employed the procedure proposed by Habiddin & Page [7] with six steps, as explained in Section 3 above. In the first step (mapping concept), it was found that differentiating electrolyte solution and non-electrolyte solution based on its electrical conductivity is the essential concept for a secondary school in Indonesia. The essential concept covers three indicators of competencies, including [1] identifying the electrical conductivity of the solution of an ionic compound, [2] identifying the electrical conductivity of the solution of covalent compound, [3] identifying the electrical conductivity of the solution of the polar covalent compound.

Next, several 22 MCQ-FR questions were constructed and intended to measure students' unscientific understanding regarding the three indicators. The example of a question in the MCQ-FR is presented in **Figure 9**. The questions were assessed in term of the scope of chemistry content and clarity in the language before being used for data collection by the chemistry lecturer and school teacher. The suggestions and feedbacks obtained were the basis for improving or revising the MCQ-FR.

In this study, the questions were focused on the conceptual type of question and avoided the algorithmic type. The initial data collection was carried out and involved five groups of students (153 in total) from two public secondary schools in Malang, East Java, Indonesia. Two groups from SMA Negeri 3 Malang (Public secondary school 3 in Malang) and three groups from SMA Negeri 8 Malang (Public secondary school 8) had taken the subject of electrolyte and non-electrolyte solutions.

Students' responses to the MCQ-FR of electrolyte and non-electrolyte solutions were categorized into scientific responses, unscientific responses and random responses. The unscientific responses were the basis to produce the FTI-ENES with 13 questions that experienced content validity afterwards. Next, the FTI-ENES was validated empirically involving two groups of students (62 in total) from SMAN 2 Ponorogo, East Java, Indonesia (Public secondary school 2 in Ponorogo). The parameters used in the empirical validation, including reliability, validity, difficulty level, discriminatory index and distractor effectiveness. Based on these parameters' values, improvements/revisions were made to refine the FTI-ENES and produce the final version of FTI-ENES.

At the same concentrations (1M), the following solutions having the strongest electrical conductivity is....

A.  $\text{HCl}(aq)$       B.  $\text{KOH}(aq)$       C.  $\text{H}_2\text{SO}_4(aq)$       D.  $\text{NaOH}(aq)$

State your reason:

.....

**Figure 9.**

*Example of a four-tier instrument in electrolyte and non-electrolyte solution.*

5.2 Results and discussion

5.2.1 Revealing students’ unscientific understanding in the topic of electrolyte and non-electrolyte solution

In the initial data collection, several students’ unscientific understanding were uncovered using the MCQ-FR. Some examples of students’ unscientific understanding that  $C_{12}H_{22}O_{11}(aq)$  is electrically conductive, partially ionized in water, and contains hydrogen bonding. Those unscientific understanding then adopted as the reason tier in the FTI-ENES, as shown in **Figure 10**.

5.2.2 The empirical validity of the FTI-ENES

The quality of the FTI-ENES is primarily reflected based on the values of 2 parameters, including validity and reliability. The two parameters are the most valuable aspect in assessing the quality of a question [31]. The last three parameters, including difficulty level, discriminatory index, and distractor effectiveness, are also essential, particularly formative and summative tests.

5.2.2.1 Validity

All the questions of the FTI-ENES instrument are valid with high validity indices. The average validity index for A-tier, R-tier and B-tier are 0.46, 0.45 and 0.53, respectively. These values confirm that the FTI-ENES is powerful for identifying students’ unscientific understanding in the area of electrolyte and non-electrolyte solutions. The detail values for each question and each tier are provided in **Table 2**.

<b>First-tier</b> Sucrose with molecular formula $C_{12}H_{22}O_{11}$ is dissolved in water. Based on its electrical conductivity, the solution can be categorized as ... A. Strong electrolyte      B. weak electrolyte      C. Non-electrolyte
<b>Second-tier</b> State the confidence rating of your answer 1. very unconfident/guessing      4. Confident 2. Unconfident      5. Absolutely confident 3. Moderate
<b>Third-tier</b> What is the reason for your answer? A. $C_{12}H_{22}O_{11}(aq)$ is not an electrically conductive solution because it does not ionize in water. B. $C_{12}H_{22}O_{11}(aq)$ ionize partially in the water. C. $C_{12}H_{22}O_{11}(aq)$ produce the mobile ions in the solution. D. $C_{12}H_{22}O_{11}(aq)$ is an electrically conductive solution because it ionizes entirely in water.
<b>Fourth-tier</b> State the confidence rating of your reason 1. very unconfident/guessing      4. Confident 2. Unconfident      5. Absolutely confident 3. Moderate

**Figure 10.**  
Example of a four-tier instrument in the FTI-ENES.



Question	Answer-tier (A-tier)		Reason-tier (R tier)		Both tier (B tier)	
	r	Category	r	Category	r	Category
1.	0.696	Valid	0.500	Valid	0.611	Valid
2.	0.495	Valid	0.372	Valid	0.564	Valid
3.	0.469	Valid	0.523	Valid	0.532	Valid
4.	0.524	Valid	0.404	Valid	0.644	Valid
5.	0.506	Valid	0.451	Valid	0.459	Valid
6.	0.455	Valid	0.485	Valid	0.515	Valid
7.	0.407	Valid	0.496	Valid	0.582	Valid
8.	0.339	Valid	0.357	Valid	0.455	Valid
9.	0.456	Valid	0.522	Valid	0.473	Valid
10.	0.592	Valid	0.583	Valid	0.697	Valid
11.	0.346	Valid	0.252	Valid	0.366	Valid
12.	0.265	Valid	0.513	Valid	0.514	Valid
13.	0.453	Valid	0.317	Valid	0.422	Valid

**Table 2.**  
*Validity indices of the FTI-ENES.*

5.2.2.2 Reliability

The reliability index of the FTI-ENES was measured using the technic of Cronbach’s Alpha. The reliability indices for A-tier, R-tier and B-tier are 0.69, 0.66 and 0.78, respectively. The values demonstrate that the instrument will produce a consistent result when it is employed over time.

5.2.2.3 Difficulty level

The difficulty level index (P) ranges from 0 to 1 and represent the number of students answering the question correctly. The higher the difficulty level value, the higher the number of students answering the question correctly, and vice versa. **Table 3** shows that the “moderate” category is the majority incident regarding the question’s difficulty level. On average, the P values for A-tier, R-tier and B-tier are 0.58, 0.53 and 0.42, respectively and fall in the “moderate” category. These values imply that the level of the questions is relevant for secondary school students.

5.2.2.4 Discriminatory index

Discriminatory index (D) compares the number of students answering the questions correctly between high achievement students and low achievement ones. The higher the D indices, the higher the number of students answering the question correctly from high achievement students, and vice versa (**Table 4**).

On average, the D values for A-tier, R-tier and B-tier are 0.53, 0.52 and 0.62, respectively and fall in the “moderate” category. These values imply that the instrument can differentiate students with high achievement and those with low achievement.

Question	Answer-tier (A-tier)		Reason-tier (R-tier)		Both tier (B-tier)	
	P	Category	P	Category	P	Category
1.	0.726	Easy	0.790	Easy	0.613	Moderate
2.	0.742	Easy	0.565	Moderate	0.484	Moderate
3.	0.516	Moderate	0.484	Moderate	0.435	Moderate
4.	0.387	Moderate	0.435	Moderate	0.274	Difficult
5.	0.581	Moderate	0.452	Moderate	0.419	Moderate
6.	0.597	Moderate	0.339	Moderate	0.323	Moderate
7.	0.677	Moderate	0.677	Moderate	0.645	Moderate
8.	0.532	Moderate	0.419	Moderate	0.226	Difficult
9.	0.468	Moderate	0.452	Moderate	0.435	Moderate
10.	0.710	Easy	0.629	Moderate	0.548	Moderate
11.	0.355	Moderate	0.726	Easy	0.306	Difficult
12.	0.726	Easy	0.565	Moderate	0.484	Moderate
13.	0.565	Moderate	0.306	Moderate	0.274	Difficult

**Table 3.**  
*The difficulty level of questions of the FTI-ENES.*

Question	A-tier		R-tier		B-tier	
	D	Category	D	Category	D	Category
1.	0.765	Excellent	0.588	Good	0.647	Good
2.	0.647	Good	0.529	Good	0.706	Good
3.	0.647	Good	0.706	Good	0.706	Good
4.	0.647	Good	0.412	Good	0.706	Good
5.	0.529	Good	0.529	Good	0.529	Good
6.	0.471	Good	0.471	Good	0.588	Good
7.	0.529	Good	0.588	Good	0.765	Excellent
8.	0.353	Good	0.471	Good	0.412	Good
9.	0.588	Good	0.647	Good	0.471	Good
10.	0.647	Good	0.765	Excellent	0.882	Excellent
11.	0.353	Good	0.176	Moderate	0.471	Good
12.	0.118	Moderate	0.647	Good	0.647	Good
13.	0.588	Good	0.176	Moderate	0.526	Good

**Table 4.**  
*Discriminatory indices of questions of the FTI-ENES.*

5.2.2.5 Distractor effectiveness

The distractor effectiveness parameter represents whether each wrong option in the A and R tiers is functional. An option is considered functional when it is chosen by at least one student [32]. **Table 5** demonstrates that all the options are functional, implying the homogeneity of the options.

Question	A tier (%)				R tier (%)			
	A	B	C	D	A	B	C	D
1.	16.13	11.29	72.58		72.58	6.45	8.06	12.90
2.	11.29	74.19	14.52		12.90	22.58	8.06	56.45
3.	33.87	14.52	51.61		17.74	48.39	6.45	27.42
4.	38.71	45.16	16.13		32.26	19.35	41.94	6.45
5.	29.03	58.06	12.90		45.16	27.42	6.45	20.97
6.	59.68	17.74	22.58		20.97	19.35	25.81	33.87
7.	11.29	67.74	11.29	9.68	17.74	9.68	64.52	8.06
8.	30.65	6.45	52.23	9.68	33.87	41.94	17.74	6.45
9.	14.52	46.77	27.42	11.29	22.58	16.13	16.13	45.16
10.	11.29	12.90	6.45	69.35	12.90	14.52	62.90	9.68
11.	25.81	24.19	35.48	14.52	46.77	25.81	19.35	8.06
12.	70.97	29.03			19.35	56.45	8.06	16.13
13.	11.29	48.39	33.87	6.45	14.52	32.26	20.97	32.26

**Table 5.**  
*Distractor effectiveness for each option each question of the FTI-ENES.*

6. Conclusions

A two-tier instrument that was initially developed by Treagust [15] is the pioneer of a multi-tier instrument. The next generation of multi-tier instruments, including three-tier, four-tier, and five-tier, responds to the drawbacks of the two-tier, which is the inability to distinguish an actual unscientific understanding and the role of guessing. We also believe that an additional drawing tier, as shown by the work of Anam et al. [28], is a rational exercise to be applied in future assessment purposes. By adopting the procedure of two-tier development, we suggest a more straightforward procedure to develop a multi-tier instrument including Mapping concept, Developing the multiple-choice question with free responses (MCQ-FR), Validating the MCQ-FR, Testing and Collecting Students’ Unscientific Understanding, Developing the prototype multi-tier instrument, and Validating the Prototype and refining the final multi-tier instrument. A wrong answer-wrong reason (WAWR) combination accompanied by a high confidence rating index (CRI) is the parameter to justify students’ unscientific understanding level. In this paper, we suggest employing a five scale CRI instead of 6 because it provides a better clarity of students to express his/her level of confidence. We also suggest that using a CRI of 3 as a limit between *genuine* and *spurious* unscientific understanding will ensure a robust justification regarding students’ unscientific understanding and lack of knowledge.

The FTI-ENES instrument developed in this study consists of 13 questions covering the topic of electrolyte and non-electrolyte solutions. The instrument’s validity and reliability revealed that it is applicable to be used in identifying students’ understanding of electrolyte and non-electrolyte solution. Even though the scope of the concepts covered in this study is relevant for secondary chemistry school, it may also be transferable for fresh university students, particularly to identify their basic chemistry knowledge gained from their learning experiences in their secondary school chemistry. Other detailed examples of the application of this procedure in developing multi-tier instruments can be found in our previous works, including in chemical kinetics [7], acid–base properties of salt solution [33, 34], and thermochemistry [30].

## Acknowledgements

We thank Directorate General Higher Education (Direktorat Jendral Pendidikan Tinggi, DIKTI), the Republic of Indonesia, for providing my PhD scholarship that contributes primarily to my multi-tier instrument project. We also thank Universitas Negeri Malang for providing a research grant through PNBP UM Scheme after finishing the PhD to continue my work in the multi-tier instrument area.

## Conflict of interest


We can declare that there is no 'conflict of interest' in this paper.

## Author details

Habiddin Habiddin\* and Septiana Ayuningrum Nofinadya  
Department of Chemistry, Universitas Negeri Malang (State University of Malang),  
Malang, Indonesia

\*Address all correspondence to: [habiddin\\_wuni@um.ac.id](mailto:habiddin_wuni@um.ac.id)

## IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 



## References

- [1] Hailikari TK, Nevgi A. How to Diagnose At-risk Students in Chemistry: The case of prior knowledge assessment. *Int J Sci Educ.* 2010;32(15):2079-2095.
- [2] Seery MK. The role of prior knowledge and student aptitude in undergraduate performance in chemistry: a correlation-prediction study. *Chem Educ Res Pract.* 2009;10(3):227-232.
- [3] Novak JD. Concept Mapping - A Useful Tool For Science-Education. *J Res Sci Teach.* 1990;27(10):937-949.
- [4] Osborne R., Gilbert JK. A Method for Investigating Concept Understanding in Science. *Eur J Sci Educ.* 1980;2(3).
- [5] Peterson RF, Treagust DF, Garnett P. Development and application of a diagnostic instrument to evaluate grade-11 and -12 students' concepts of covalent bonding and structure following a course of instruction. *J Res Sci Teach.* 1989;26(4):301-314.
- [6] Taber KS. Ideas About Ionisation Energy: A Diagnostic Instrument. *Sch Sci Rev.* 1999;81(295):97-104.
- [7] Habiddin H, Page EM. Development and validation of a four-tier diagnostic instrument for chemical kinetics (FTDICK). *Indones J Chem.* 2019;19(3):720-736.
- [8] Kinchin IM. Using Concept Maps to Reveal Understanding: A Two-Tier Analysis. *Sch Sci Rev.* 2000;81(296):41-46.
- [9] Chandrasegaran AL, Treagust DF, Mocerino M. The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chem Educ Res Pract.* 2007;8(3):293-307.
- [10] Towns MH, Robinson WR. Student Use Of Test-Wiseness Strategies In Solving Multiple-Choice Chemistry Examinations. *J Res Sci Teach.* 1993;30(7):709-722.
- [11] Pesman H, Eryilmaz A. Development of a Three-Tier Test to Assess Misconceptions About Simple Electric Circuits. *J Educ Res.* 2010;103(3):208-222.
- [12] Dindar AC, Geban O. Development of a three-tier test to assess high school students' understanding of acids and bases. In: 3rd World Conference on Educational Sciences. 2011.
- [13] Habiddin H, Akbar DFK, Aziz AN, Hasan H, Mustapa K. Developing a multi-tier instrument for chemistry teaching: A challenging exercise. *AIP Conf Proc.* 2021;2330(1):20001.
- [14] Kilinc A, Yeşiltaş NK, Kartal T, Demiral Ü, Eroğlu B. School Students' Conceptions about Biodiversity Loss: Definitions, Reasons, Results and Solutions. *Res Sci Educ.* 2013;43(6):2277-2307.
- [15] Treagust DF. Development and use of diagnostic tests to evaluate students' misconceptions in science. *Int J Sci Educ.* 1988;10(2):159-169.
- [16] Tan K-CD, Goh NK, Chia LS, Treagust DF. Development and application of a two-tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. *J Res Sci Teach.* 2002;39(4):283-301.
- [17] Tuysuz C. Development of two-tier diagnostic instrument and assess students' understanding in chemistry. *Sci Res Essays.* 2009;4(6):626-631.
- [18] Griffard PH, Wandersee JH. The two-tier instrument on photosynthesis:

What does it diagnose? *Int J Sci Educ.* 2010;23(10):1039-1052.

[19] Tyson L, Treagust DF, Bucat RB. The complexity of teaching and learning chemical equilibrium. *J Chem Educ.* 1999;76(4):554-558.

[20] Adadan E, Savasci F. An analysis of 16-17-year-old students' understanding of solution chemistry concepts using a two-tier diagnostic instrument. *Int J Sci Educ.* 2012;34(4):513-544.

[21] Caleon I, Subramaniam R. Do Students Know What They Know and What They Don't Know? Using a Four-Tier Diagnostic Test to Assess the Nature of Students' Alternative Conceptions. *Res Sci Educ.* 2010;40(3):313-337.

[22] Hasan S, Bagayoko D, Kelley EL. Misconceptions and the Certainty of Response Index (CRI). *Phys Educ.* 1999;34(5):294-299.

[23] Arslan HO, Cigdemoglu C, Moseley C. A Three-Tier Diagnostic Test to Assess Pre-Service Teachers' Misconceptions about Global Warming, Greenhouse Effect, Ozone Layer Depletion, and Acid Rain. *Int J Sci Educ.* 2012;34(11):1667-1686.

[24] Kirbulut ZD. Using Three-Tier Diagnostic Test to Assess Students' Misconceptions of States of Matter. *EURASIA J Math Sci Technol Educ.* 2014;10(5):509-521.

[25] Cetin-Dindar A, Geban O. Development of a three-tier test to assess high school students' understanding of acids and bases. *Procedia - Soc Behav Sci.* 2011;15:600-604.

[26] Aydeniz M, Bilican K, Kirbulut ZD. Exploring Pre-Service Elementary Science Teachers' Conceptual Understanding of Particulate Nature of Matter through Three-Tier Diagnostic Test. *Int J Educ Math.* 2017;5(3): 221-234.

[27] Sreenivasulu B, Subramaniam R. University Students' Understanding of Chemical Thermodynamics. *Int J Sci Educ.* 2013;35(4):601-635.

[28] Anam RS, Widodo A, Sopandi W, Wu HK. Developing a five-tier diagnostic test to identify students' misconceptions in science: an example of the heat transfer concepts. *Elem Educ Online.* 2019;18(3):1014-1029.

[29] Orlich DC, Harder RJ, Callahan RC, Trevisan MS, Brown AH. *Teaching Strategies: A Guide to Effective Instruction.* 9th ed. Boston: Wadsworth Publishing; 2010.

[30] Habiddin H, Utari JL, Muarifin M. Diagnostic tool to reveal students' conception on thermochemistry. *AIP Conf Proc.* 2021;2330(1):20008.

[31] Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments used in research. *Am J Heal Syst Pharm.* 2008/11/21. 2008;65(23): 2276-84.

[32] DiBattista D, Kurzawa L. Examination of the Quality of Multiple-choice Items on Classroom Tests. *Can J Scholarsh Teach Learn CJSOTL.* 2011;2(2).

[33] Habiddin H, Ameliana DN, Sua'idy M. Development of a Four-Tier Instrument of Acid-Base properties of salt Solution. *JCER (Journal Chem Educ Res.* 2020;4(1):51-7.

[34] Husniah I, Habiddin H, Sua'idy M, Nuryono N. Validating an instrument to investigate students' conception of Salt hydrolysis. *J Disruptive Learn Innov.* 2019;1(1):1-6.